Journal of Insect Science
www.insectscience.org

Expressed sequence tags from cephalic chemosensory organs of the northern walnut husk fly, *Rhagoletis suavis*, including a putative canonical odorant receptor

Karlene M. M. Ramsdell[1a], Sheila A. Lyons-Sobaski[2b], Hugh M. Robertson[1c], Kimberly K. O. Walden[1d], Jeffrey L. Feder[3e], Kevin Wanner[1f], and Stewart H. Berlocher[1g]

[1]Department of Entomology, 320 Morrill Hall, 505 S. Goodwin Ave., University of Illinois at Urbana-Champaign, Urbana, IL 61801
[2]Department of Biology, Albion College, 611 E. Porter St., Albion, Michigan 49224
[3]Department of Biological Sciences, P.O. Box 369, Galvin Life Science Center, University of Notre Dame, Notre Dame, IN 46556-0369.

Abstract

Rhagoletis fruit flies are important both as major agricultural pests and as model organisms for the study of adaptation to new host plants and host race formation. Response to fruit odor plays a critical role in such adaptation. To better understand olfaction in *Rhagoletis*, an expressed sequence tag (EST) study was carried out on the antennae and maxillary palps of *Rhagoletis suavis* (Loew) (Diptera: Tephritidae), a common pest of walnuts in eastern United States. After cDNA cloning and sequencing, 544 ESTs were annotated. Of these, 66% had an open reading frame and could be matched to a previously sequenced gene. Based on BLAST sequence homology, 9% (49 of 544 sequences) were nuclear genes potentially involved in olfaction. The most significant finding is a putative odorant receptor (OR), *RSOr1*, that is homologous to *Drosophila melanogaster Or49a* and *Or85f*. This is the first tephritid OR discovered that might recognize a specific odorant. Other olfactory genes recovered included odorant binding proteins, chemosensory proteins, and putative odorant degrading enzymes.

Keywords: host race, *Juglans nigra*, olfaction, odorant receptor, *Rhagoletis*, Tephritidae, speciation
Abbreviations: CSP, chemosensory protein; **EST,** expressed sequence tag; **OBP,** odorant binding protein; **OR,** odorant receptor
Correspondence: [a] ramsdell@life.uiuc.edu, [b] ssobaski@albion.edu, [c] hughrobe@uiuc.edu, [d] k-walden@life.uiuc.edu, [e] jfeder@nd.edu, [f] kwanner@life.uiuc.edu, [g] stewartb@life.uiuc.edu
Associate Editor: Zhijian Tu was editor of this paper.
Received: 9 June 2008, **Accepted:** 7 September 2008

ISSN: 1536-2442 | Vol. 10, Number 51

Cite this paper as:
Ramsdell KMM, Lyons-Sobaski SA, Robertson HM, Walden KKO, Feder JL, Wanner K, Berlocher SH. 2010. Expressed sequence tags from cephalic chemosensory organs of the northern walnut husk fly, *Rhagoletis suavis*, including a putative canonical odorant receptor. 11 pp. *Journal of Insect Science* 10:51 available online: insectsicence.org/10.51

Introduction

Species of the genus *Rhagoletis* are important pests of fruits such as apples, cherries, tomatoes, walnuts, and blueberries. They are equally important as the focus of the debate about the possibility of sympatric speciation via the formation of host races on new host plants (Bush 1966; Berlocher and Feder 2002). In the case of the apple host race of *Rhagoletis pomonella*, two key adaptations arose approximately 150 years ago in the ancestral (and still extant) hawthorn race that allowed colonization of apple. One is alteration of the olfactory response so that both sexes are attracted to the odor of the new host apple (Linn et al. 2003, 2004; Dambroski et al. 2005), and the other is shifting life history phenology to match the fruit ripening time of apple (Filchack et al. 2000). This study is the first attempt to catalog genes involved in olfaction in *Rhagoletis* by carrying out an expressed sequence tag (EST) project on the antennae and maxillary palps of *Rhagoletis suavis* (Loew) (Diptera: Tephritidae). This species was used because it can be obtained more easily in the large numbers required for an EST project on olfactory organs than *R. pomonella* can.

Many features of the molecular biology of olfaction in *Rhagoletis* can be anticipated from what is known of olfaction in *Drosophila melanogaster*, which is a key model organism for studying olfaction (Rützler and Zwiebel 2005; Hallem et al. 2006; Vosshall and Stocker 2007). Two major gene families involved are the odorant binding proteins (OBPs) (Hekmat-Scafe et al. 2002) and the odorant receptors (ORs) (Robertson et al. 2003). OBPs are usually highly expressed, which makes detection in antennal EST projects likely (e.g. Robertson et al.1999); whereas ORs are generally expressed at such low levels that they are difficult to obtain with this method. Based on the *D. melanogaster* genome, it was anticipated that the most important recoveries from this EST project would be key olfactory gene products such as ORs and OBPs. However, other classes of genes have been proposed as having a possible role in olfaction, such as chemosensory proteins (Briand et al. 2002; Lartigue et al. 2002) and odorant degrading enzymes, as well as genes that are of general interest.

Materials and Methods

Flies and collection of antennae and palps

Collection of large numbers of *Rhagoletis* flies is most easily accomplished by rearing larvae from infested fruit (*Rhagoletis* life history is described by Boller and Prokopy 1975). In the fall of 2000, approximately 50,000 *R. suavis* (Loew) larvae were reared from black walnut, *Juglans nigra* L. (Fagales: Juglandaceae), fruit from sites near White Heath, Illinois (Piatt County). Pupae were placed in a 4° C cold room to break diapause and then removed in batches throughout the spring of 2001. Emerging flies were placed in cages with food and water (Prokopy and Bush 1973) until they could be processed. Processing was carried out as rapidly as possible after eclosion because young adults were assumed to have the highest expression of olfactory receptors. Heads from live flies were removed and accumulated at -80° C. The day before RNA extraction, the frozen heads were shaken on a soil sieve to harvest antennae. Maxillary palps and major head bristles that may also have chemoreceptors were harvested incidentally. Maxillary palps have sensilla used in odor recognition and express gustatory receptors, but mRNA would

not have been obtained from major bristles because the cell bodies are not in the bristles. The shaking and sieving was not severe enough to break the heads, so there was no contamination from brain or eye tissues.

RNA extraction and cDNA library construction

Total RNA was isolated from antennae and maxillary palps using a guanidinium thiocyanate/phenol-chloroform extraction protocol (RNA Isolation Kit, Stratagene, www.stratagene.com). mRNA was purified from total RNA using a Poly(A) Quik® mRNA Isolation Kit (Stratagene, www.stratagene.com), which utilizes an oligo-dT cellulose column. A unidirectional plasmid cytomegalovirus-polymerase chain reaction cDNA library primed with oligo-dT was constructed by Stratagene using PCR amplification. The plasmid library was transformed into Stratagene's host strain Epicurian Coli® XL-10 Gold™. For further details of molecular methods and results see Ramsdell (2004).

Clone sampling & DNA sequencing

Plasmid clones were sampled by plating the library onto LB-kanamycin agar and picking colonies. Colonies were individually transferred to 96-well plates. Each well contained 80 µl of a 30% (v/v) glycerol-LB mixture. Six plates were prepared and submitted to the W.M. Keck Center for Comparative and Functional Genomics (University of Illinois at Urbana-Champaign) for sequencing from the 5' end using ABI automation. Clones of interest were cultured and purified, and the insert was sequenced from both directions when necessary.

Sequence analysis

Sequences were edited with Microsoft Excel® and BBEdit Lite (Bare Bones Software, Inc., www.barebones.com). DNA Strider v1.1 (Marck 1988) was used for protein translations, to find open reading frames, convert reverse complement sequence reads, and generate Kyte-Doolittle hydropathy plots (Kyte and Doolittle1982). BLAST (Altschul et al.,1990, 1997) was used with networked servers (National Center for Biotechnology Information; http://www.ncbi.nlm.nih.gov) to find the most similar sequence matches to the *R. suavis* ESTs in the GenBank databases. Significantly similar matches had E values of 10^{-4}. An initial screen using the tblastn option (translated DNA query searching translated DNA database) was followed with both nucleotide and protein BLAST searches (blastn, blastx, blastp) of the largest open reading frames. Searches were generally restricted to Diptera. Sequences of interest were aligned using Clustal X 2.0 (Larkin et al. 2007) using default settings. EST sequences were deposited in the dbEST EST database at the National Center for Biotechnology Information (Accessions EX453814 - EX454354).

To show the relationships of the *Drosophila melanogaster* OR sequences that are most similar to the *R. suavis* OR, a neighbor-joining tree of corrected distances was built using Clustal X (Larkin et al. 2007). Bootstrapping was performed with Clustal X with 10,000 pseudoreplications.

Results

Recovery of ESTs

A total of 544 clones was sequenced, with an average length 532.02 ± SE 9.88 bp (range 14 to 967 bb). A wide variety of gene transcripts was obtained. As expected from a normalized library, 418 (76.8%) of the sequences were unique. The largest number of duplicates was

18 for a sequence similar to DmCG13095 (a peptidase). Of the 544 total sequences, 186 had no obvious ORF and did not produce a significant BLASTx match with a known protein sequence in GenBank. Of the 358 sequences with an ORF, 86 produced either a weak match with a known gene, or a low E-value match with a sequence of unknown function. Of the 313 sequences with a significant match to a sequence with a known function, 37 were mitochondrial and 276 were nuclear. As expected, protein BLAST searches yielded much smaller E values than did nucleotide searches. The exceptions involved nucleotide matches with sequences from other tephritid flies (*R. pomonella* and the medfly *Ceratitis capitata*), which usually resulted in the smallest E values, presumably because 3' UTRs retained some sequence similarity.

A representative set of the nuclear matches is shown in 3 1. Given that mRNA was extracted from antennae and maxillary palps, it is not surprising that 48 (9%) of the sequences had a function or putative function relating to chemoreception. Also, 24 of the sequences had a known or possible role in development. This finding is not surprising because the source flies were young adults that were not fully mature. Also included in Table 1 are a few sequences that have been implicated in diapause and life history; such genes were not the target of this study, but they are noted because diapause is critical to host race formation in *Rhagoletis*. Although they do not appear to play a role in diapause initiation, heat shock loci can be up-regulated during diapause (Rinehart et al. 2007).

Chemosensory proteins

Thirteen sequences were recovered that coded for two different chemosensory proteins (CSPs), RsCSP1 and CSP2. The *R. suavis* CSPs matched only chemosensory proteins in the public databases and were identified as belonging to the conserved domain of the CSP family. Proteins from *D. melanogaster* Antennal Protein 10 (A10 or OS-D) and Ejaculatory Bulb Protein III (PEBme III), were the best matches for RsCSP1 and RsCSP2, respectively. A10 and RsCSP1 had a pairwise amino acid identity of 66%, and RsCSP2 and PEBme III were 82% identical. The *R. suavis* CSPs have an amino acid identity of 45.7%; the mature forms are 50.9% identical. RsCSP1 was 155 amino acids in length, including a signal peptide of 21 amino acids, and RsCSP2 had a length of 127 with its 18 amino acid signal.

Odorant binding proteins

Nine OBPs, RsObp1 to RsObp9, were recovered. All had top matches to dipteran OBPs in the public databases. The Kyte-Doolittle hydropathy plots of the nine proteins showed typical OBP profiles with hydrophobic peptide signals (Peng and Leal 2001). Including their peptide signals, the OBPs ranged in length from 124 to 164 amino acids. Overall, the *R. suavis* OBPs were diverse and showed little conservation of amino acid residues. The mature OBPs had mean pairwise amino acid identities of 19.9%, with a range of 7.4 to 55.9%. Signal peptides were 15 to 26 amino acids in length (Ramsdell 2004).

Odorant receptor protein

The *R. suavis* OR sequence (EX453813, 634 bp) was identified as an OR because a protein BLAST search of a 450 bp/150 amino acid ORF significantly (2E-04) matched *DmOr49a*. Resequencing of the clone from both ends revealed an unambiguous match with two *D. melanogaster* OR sequences. These were *DmOr49a* (4E-56, amino acid

Table 1. Genes of interest recovered in *R. suavis* antennal EST study.

Broad category	Specific type	Symbol	dbEST Accession	BP	Annotation ID	species	Locus	Description	E value
Chemoreception	receptors	RsOr1	EX453813 / EU204908	1290*	CG13158	*D. melanogaster*	Or49a	Odorant receptor 49a. Almost complete sequence is EU204908. Chemoreception.	1E-52*
	odorant binding proteins	RsObp1	EX453814 - EX453819	758	CG10436	*D. melanogaster*	Pbprp1	Pheromone-binding protein-related protein 1. Chemoreception.	7.00E-24
		RsObp2	EX453820	679	CG11421	*D. melanogaster*	Pbprp3	Pheromone-binding protein-related protein 3. Chemoreception.	1.00E-41
		RsObp3	EX453821	777	CG11421	*D. melanogaster*	Pbprp3	Pheromone-binding protein-related protein 3. Chemoreception.	2.00E-56
		RsObp4	EX453822 - EX453823	736	CG1176	*D. melanogaster*	Pbprp4	Pheromone-binding protein-related protein 4. Chemoreception.	3.00E-31
		RsObp5	EX453824	659	CG12665	*D. melanogaster*	Obp8a	Odorant-binding protein 8a. Chemoreception, olfactory behavior.	8.00E-08
		RsObp6	EX453825 - EX453826	613	CG13873	*D. melanogaster*	Obp56g	Odorant-binding protein 56g. Chemoreception.	4.00E-11
		RsObp7	EX453827	652	CG1668	*D. melanogaster*	Pbprp2	Pheromone-binding protein-related protein 2 Obp19d. Chemoreception.	1.00E-27
		RsObp8	EX453828 - EX453830	707	CG6641	*D. melanogaster*	Pbprp5	Pheromone-binding protein-related protein 5. Chemoreception.	2.00E-36
		RsObp9	EX453831 - EX453833	644	CG6641	*D. melanogaster*	Pbprp5	*R. suavis* paralog of Pbpr5. Chemoreception.	4.00E-21
	chemosensory proteins	RsCSP1	EX453848 - EX453857	583	CG6642	*D. melanogaster*	a10	Antennal protein 10. Expressed in CNS and male sex organs as well as antennae.	1.00E-46
		RsCSP1	EX453844-EX453847	753	CG11390	*D. melanogaster*	PebIII	Ejaculatory Bulb Protein III (PEBme III)	6.00E-58
	odorant degrading		EX453836 - EX453837	716	CG17639	*D. melanogaster*	CG17639	Glutathione-S-transferase. Detoxification.	1.00E-97
			EX453839 - EX453840	771	CG6649	*D. melanogaster*	Ugt35b	Antennal-enriched UDP-glycosyltransferase. Detoxification.	1.00E-57
	other		EX453861	762	CG5430	*D. melanogaster*	a5	Antennal protein 5. Signal transduction.	9.00E-51
Life history/diapause	circadian rythym		EX453862	626	AY788366	*Ceratitis capitata*	per	Period. Circadian rythym regulation. Similar to *D. melanogaster* G2647.	6.00E-29
	heat shock proteins		EX453872 - EX453873	661	AM084221	*Ceratitis capitata*	hsp83	Heat shock protein 83. Defense. Similar to *D. melanogaster* CG1242.	1.00E-117
			EX453874 - EX453875	772	EF103584	*R. pomonella*	hsp68	Heat shock protein 68. Defense. EF103584 reported as hsp70, but is slightly more similar to *D. melanogaster* hsp68 (CG5436).	1E-153‡
			EX453876 - EX453877	587	AAC23392	*Ceratitis capitata*	Hsp70	Heat shock protein 70 family. Defense. Similiar to *D. melanogaster* Hsp70-4 (CG4264).	2E-49‡
			EX453878	696	CG8863	*D. melanogaster*	CG8863	Heat shock protein 40 binding protein.	1.00E-76
Development			EX453903	522	CG11785	*D. melanogaster*	bai	Baiser. Determination of dorsal/ventral axis.	1.00E-38
			EX453905	541	CG1780	*D. melanogaster*	Idgf4	Imaginal disc growth factor 4.	2.00E-65
			EX453906	565	CG2210	*D. melanogaster*	awd	Abnormal wing discs.	1.00E-80
			EX453908	678	CG3260	*D. melanogaster*	Zfrp8	Zinc finger protein RP-8. Early embryonic development.	6.00E-48
			EX453913	596	CG4254	*D. melanogaster*	tsr	Twinstar. Axon growth regulation.	5.00E-93
			EX453914	555	CG7762	*D. melanogaster*	Rpn1	Ribophorin I. Cell cycle regulation.	7.00E-89
			EX453915	611	CG8440	*D. melanogaster*	Lis-1	Lissencephaly-1. Anatomical development.	E-127
			EX453916	588	CG8567	*D. melanogaster*	Deaf1	Deformed epidermal autoregulatory factor-1. Early embryonic development.	5.00E-19
			EX453924	341	CG9635	*D. melanogaster*	RhoGEF2	Rho guanyl-nucleotide exchange factor 2. Anatomical devvelopment.	7.00E-19
Gene regulation			EX453925	764	CG3644	*D. melanogaster*	bic	Bicaudal. Transcription factor; RNA pol II regulation.	2.00E-55
Pheromone production			EX454031 - EX454032	776	CG9747	*D. melanogaster*	CG9747	Acyl-CoA delta11-desaturase. Fatty acid biosynthesis. Also similar to desat2 (CG5925).	2.00E-37
Marker loci	allozyme		EX454045	763	CG11793	*D. melanogaster*	Sod	Superoxide dismutase.	6.00E-83
			EX454047	772	CG3140	*D. melanogaster*	Adk2	Adenylate kinase-2.	E-128
			EX454048	563	CG32031	*D. melanogaster*	Argk	Arginine kinase.	2.00E-73
			EX454049	780	CG6058	*D. melanogaster*	Ald	Aldolase.	E-148
			EX454050	604	CG9042	*D. melanogaster*	G-3-pdh	Glycerol-3-P dehydrogenase.	3.00E-77
	genomic sequence		EX454051	397	AY930988	*R. pomonella*	P3072	*Rhagoletis* EST locus.	7E-55‡
			EX454052	359	AY931003	*R. electromorpha*	P3072	*Rhagoletis* EST locus.	1E-52‡

*Length and probability are for the almost complete sequence EU204908, not the EST RSUA001. ‡Probabilities are for nucleotide blast, not tblastx (translated nucleotide database searched with a translated nucleotide query), as in the rest of the searches. See text.

identity = 31%) and *DmOr85f* (1E-37, amino acid identity = 26%). The alignment of these three sequences is shown in Figure 1. Based on the alignment, it is likely that a few amino acids were missing at the N-terminus of the *R. suavis* sequence. To increase the likelihood that the nearest known homolog of the *R. suavis* receptor was found, the nine *Drosophila* OR sequences were included in the neighbor-joining tree analysis, ranked in order of decreasing E value, between *DmOr49a* and an *Anopheles gambiae* receptor (AGAP001912, 8E-28). The resulting neighbor-joining tree (Figure 2, shows only the relevant part of tree, including Or85f from *Drosophila pseuodoobscura* supports the conclusion that the *D. melanogaster* homolog of the *R. suavis* odorant receptor sequence, henceforth RsOr1, was *DmOr49a*. The RsOr1 sequence clearly showed the characteristic hydropathy plot of a 7-transmembrane protein, with alternating hydrophobic and hydrophilic regions (Figure 3).

Discussion

Chemosensory Proteins

The function of CSPs is not clear at this point. They are highly expressed in insect antennae, and some work supports a role as olfactory

```
DM49a           -MEKLRSYEDFIFMANMMFKTLGYDLFHTPKPWWRYLLVRGYFVLCTISNFYEASMVTTR 59
RHAG            ---------QVFWGPNALFRAVGYDFQRLPRPYWRQILMKAVLIFMILSAICIRIYMFMS 51
DM85f           MEPVQYSYEDFARLPTTVFWIMGYDMLGVPKTRSRRILYWIYRFLCLASHGVCVGVMVFR 60
                         :.  .. :*  :***:   *:. * :*     .:   *       :

DM49a           IIEWESLAGSPSKIMRQGLHFFYMLSSQLKFITFMINRKRLLQLSHRLKELYPHKEQNQR 119
RHAG            LRELIIRDDILN-SFRLGAFIAYGVDSNVKFAYFIFKAHRLRKIYDFLAAEYPQTSSEQK 110
DM85f           MVEAKTID-NVSLIMRYATLVTYIINSDTKFATVLQRS-AIQSLNSKLAELYPKTTLDRI 118
                : *        . :* . . * :.*: **  .: .  : .:   *   **:. ::

DM49a           KYEVNKYYLSCSTRNVLYVYYFVMVVMALEPLVQSCIMYLI-----GFGKADFTYKRIFP 174
RHAG            LYKIDIYGFQRAP-VMICAYMAVVASIMLSPLLQSIVTYIIDIYRFGYDAAEYPYLHPIP 169
DM85f           YHRVNDHYWTKSFVYLVIIYIGSSIMVVIGPIITSIIAYFT--------HNVFTYMHCYP 170
                 :.:: :    :   ::  *      : : *:: * : *:            :.* :  *

DM49a           -TRLTFDSEKPLGYVLAYVIDFTYSQFIVNVSLGTDLWMMCVSSQISMHLGYLANMLASI 233
RHAG            -MPYNFDYYTPRYYIPVYMVESLNGHFSSTTNLGTDLFISIFSGQLCMQLEYLGYSLETY 228
DM85f           YFLYDPEKDPVWIYISIYALEWLHSTQMVISNIGADIWLLYFQVQINLHFRGIIRSLADH 230
                      :      *:  * ::   .     .:*:*:::  .. *: :::  :   *

DM49a           RPSPETEQQDCDFLASIIKRHQLMIRLQKDVNYVFGLLLASNLFTTSCLLCCMAYYTVVE 293
RHAG            EPSMEKSEDDCEFLRKWIRKHQLMLGLCADLDEVFGTTLLCKLITNCTYFCIIVAQLMLE 288
DM85f           KPSVKHDQEDRKFIAKIVDKQVHLVSLQNDLNGIFGKSLLLSLLTTAAVICTVAVYTLIQ 290
                .** : .::* .*: . : ::  :: *  *:: :**  *  .*:*.. :* :.  :::

DM49a           GFNWEGISYMMLFASVAAQFYVVSSHGQMLIDLSTNLAKAAFESKWYEGSLRYKKEILIL 353
RHAG            GYGYGFLNFGSFFFLTVAQFFMVCQYGQNLITISEHLSFSAYKNRWYNGSKAYKKMILTI 348
DM85f           GPTLEGFTYVIFIGTSVMQVYLVCYYGQQVLDLSGEVAHAVYNHDFHDASIAYKRYLLII 350
                *     :.:  ::   . *.::*. :** :: :* .:: :.::  :::.*  **: :* :

DM49a           MAQAQRPLEISARGVIIISLDTFKILMTITYRFFAVIRQTVEK 396
RHAG            ITRAQTPANLTAKGFQPISLLTFQIVMSVTYRVFAVLQQVFD- 390
DM85f           IIRAQQPVELNAMGYLSISLDTFKQLMSVSYRVITMLMQMIQ- 392
                : :** * ::.* *   *** **: :*:::**.:::: * .:
```

Figure 1. Clustal X alignment of the *Rhagoletis suavis* receptor RsOr1, and *Drosophila melanogaster Or49a* and *85f*, which are the most similar sequences to RsOr1. Symbols: * = identical amino acids, : = conservative substitutions, . = semiconservative substitutions. High quality figures are available online.

ligand transporters (Briand et al. 2002; Lartigue et al. 2002). Recent work in *Bombyx mori*, however, indicates that they are commonly expressed in many parts of the body in addition to antennae (Conga et al. 2007). The fact that two different CSPs were recovered in this small study of 544 ESTs indicates that, consistent with other work, CSPs are highly expressed in antennae, but their possible role in *Rhagoletis* olfaction remains uncertain.

Odorant binding proteins

Drosophila melanogaster has 51 OBPs (Hallem et al. 2006, Hallem and Carlson 2006). Thus the recovery of nine different *R. suavis* OBP sequences, all with *D. melanogaster* orthologues, (Table 1) from only 544 ESTs suggests that most, if not all, of the *R. suavis* OBPs could be recovered by a modestly more extensive EST study. The exact role that OBPs could play in host specificity remains unknown; however, it is quite likely that they play a significant part. Recent work on *Drosophila* pheromone reception demonstrates both that OBPs are necessary for chemoreception and that some are highly specific for particular odorants (Xu et al. 2005).

The *R. suavis* OBPs have a mean pairwise amino acid identity of about 20%, which is typical for phylogenetically distant members of the OBP gene family (Robertson et al. 1999). Their diversity, coupled with their apparent homology to *D. melanogaster* OBPs, make them good candidates for use as genetic tools in studies of acalypteran and other dipteran lineages.

The odorant receptor sequence

Odorant receptors are believed to play a critical role in the host finding behavior in insects, yet they are difficult to obtain without

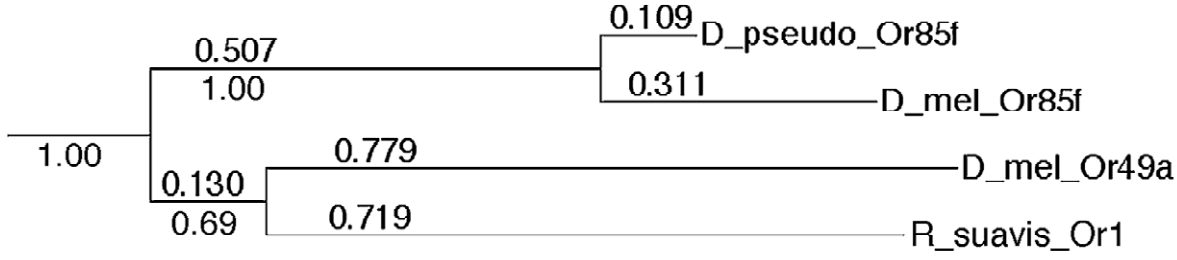

Figure 2. Neighbor-joining tree of sequences similar to the *Rhagoletis suavis* odorant receptor (*Drosophila pseudooscura* Or85f, and *D. melanogaster* Or85f and Or49a). Distances are percent dissimilarities corrected for multiple replacements. Values under the distances are bootstrap values. High quality figures are available online.

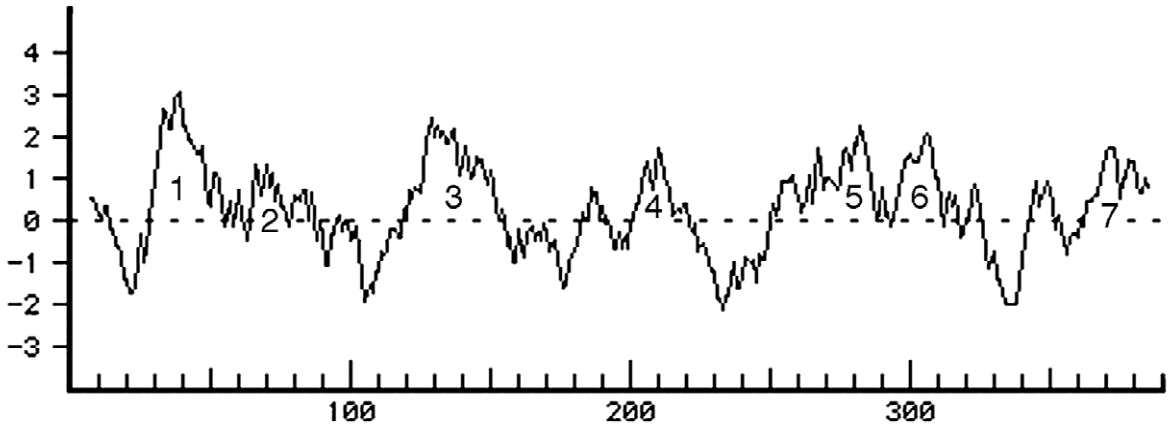

Figure 3. Kyte-Doolittle hydropathy plot for RsOr1. High quality figures are available online.

completely sequenced genomes. Only a few odorant receptor sequences have been discovered in insect EST projects (and none of these are published), suggesting a low rate of expression. Indeed Vosshall et al. (1999) noted that *D. melanogaster* ORs were present in fewer than 1 in 500,000 clones in an antennal library. This is lower than this rate of 1 OR in 544 clones, but it is likely that a *Rhagoletis* genome will be necessary to obtain a complete set of OR genes.

RsOr1 is significant as the first reported putative ligand-binding receptor from a tephritid fly. It is not the first tephritid receptor; that distinction belongs to a receptor recovered from *C. capitata* by Larrson et al. (2004). However, the *C. capitata* OR was homologous to the atypical, "non-canonical" *Or83b*, which plays a role in localizing conventional or “canonical” receptors to the membrane and is highly conserved across insects (Jones et al. 2005). But *Or83b* does not bind odorant ligands, and, thus, its homologs are unlikely to play a direct part in host plant adaptation. *RsOr1*, on the other hand, was clearly homologous to the canonical *DmOr49a*. Unfortunately, it is not possible, at this point, to speculate on the volatile, or volatiles, which elicits a response from *RsOr1*, as the ligands of *DmOr49a* have not yet been determined (Hallem and Carlson 2006).

However, it is probable that the OR sequences, or their expression patterns, or both, differ substantially between *R. suavis* and the apple maggot *R. pomonella*. The fruit volatiles of apples are characterized by high concentrations of esters (Linn et al. 2003; Souleyre et al. 2005), while those of the ancestral host of *R. pomonella*, hawthorns, are characterized by ethyl acetate, long-chain alcohols, and various aldehydes (Linn et al. 2003). But a completely different spectrum of volatiles, dominated by terpenes and terpenoids, occurs in walnut fruits (Hennemanm et al. 2002). Moreover, many of the walnut terpenoids, such as β-pinene, limonene, β-caryophylene, and α-humulene (Hennemanm et al. 2002) did not elicit any responses from the (incomplete) set of ORs tested by Hallem and Carlson (2006). Thus *R. suavis* may provide insights into insect olfaction that are not possible with *Drosophila*.

R. suavis may be a good species with which to study the various roles of odorant degrading enzymes in olfaction. As pointed out by Rützler and Zwiebel (2005), odorant degrading enzymes are necessary to remove the signaling molecule after a cell response has been initiated and also because chemosensory systems must be open to the environment. Odorant degrading enzymes may have a secondary role of degrading toxic odorants before they can cause cellular damage. One of the major components of walnut fruit odor is limonene, which is used as an insecticide, and also causes "spontaneous stimulation of sensory nerves” (Weinzierl 1998, p. 106; mechanism not known). Detoxification of limonene in the cutworm *Spodoptera* is reported to be similar to mammalian detoxification (Miyazawa et al. 1998), where oxidative degradation by cytochrome P450s appears to be the most important pathway (e.g., Miyazawa et al. 2002). No cyt P450 sequences were recovered in this study, but they represent one of several pathways that should be studied in olfaction in phytophagous insects.

While tremendous strides have been made in understanding the molecular biology of chemosensation in recent years (Rützler M and, Zwiebel LJ. 2005, Hallem et al. 2006,

Vosshall and Stocker 2007), we are still very far from being able to understand the relative importance for host adaptation of peripheral vs. central processes, sequence vs. expression differences, or even the relative importance of the different classes of genes involved. Koop et al. (2008) have recently demonstrated that expression differences for both Ors ORs and OBPs have been involved in the adaptation of *Drosophila sechellia* to its food plant *Morinda citrifolia*. But more classes of molecules will need to be included in future such studies. ODEs and CSPs will certainly need to be added. But even genes that seemingly have little to do with olfaction may be important. For example, Hsp70 genes could affect receptor function in chemosensory cells because of their role of in guiding the folding of proteins (Bukau et al. 2006).

Acknowledgments

We thank Steve Ramsdell for the many hours he spent collecting and processing black walnuts. Without his assistance, obtaining the 50,000 flies necessary for this study would not have been possible. We thank an anonymous reviewer for the suggestion that Hsp70 genes could play a role in olfaction. NSF DEB-99-77011, NSF DEB 06-14528, and AG 2007-35604-17886 provided support.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215(3): 403-410.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* 25(17): 3389-3402.

Berlocher SH, Feder JL. 2002. Sympatric speciation in phytophagous insects: Moving beyond controversy? *Annual Review of Entomology* 47: 773-815.

Boller, EF, Prokopy RJ. 1975. Bionomics and management of *Rhagoletis*. *Annual Review of Entomology* 21: 223-246.

Briand L, Nespoulous C, Huet JC, Takahashi M, Pernollet JC. 2001. Ligand binding and physico-chemical properties of ASP2, a recombinant odorant-binding protein from honeybee (*Apis mellifera* L.). *European Journal of Biochemistry* 268(3): 752-760.

Bukau B, Weissman J, Horwich A. 2006. Molecular Chaperones and Protein Quality Control. *Cell* 125: 443-451.

Bush GL. 1966. The taxonomy, cytology, and evolution of the genus *Rhagoletis* in North America (Diptera, Tephritidae). *Bulletin of the Museum of Comparative Zoology* 134: 431-562.

Dambroski HR, Linn C, Berlocher S, Forbes AA, Roelofs W, Feder JL. 2005. The genetic basis for fruit odor discrimination in *Rhagoletis* flies and its significance for sympatric host shifts. *Evolution* 59(9): 1953-1964.

Filchak KE, Roethele JB, Feder JL. 2000. Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature* 407: 739-742.

Gonga DP, Zhanga H, Zhaoa P, Lina Y, Xia QY, Xianga ZH. 2007. Identification and expression pattern of the chemosensory protein gene family in the silkworm, *Bombyx*

mori. *Insect Biochemistry and Molecular Biology* 37(3): 266-277.

Hallem EA, Carlson JR. 2006. Coding of odors by a receptor repertoire. *Cell* 125(1): 143-160.

Hallem EA, Dahanukar A, Carlson JR. 2006. Insect odor and taste receptors. *Annual Review of Entomology* 51: 113-135.

Hekmat-Scafe DS, Scafe DR, McKinney AJ, Tanouye MA. 2002. Genome-wide analysis of the odourant-binding protein gene family in *Drosophila melanogaster*. *Genome Research* 12(9): 1357-1369.

Hennemanm L, Dyreson EG, Takabayashi J, Raguso RA. 2002. Response to walnut olfactory and visual cues by the parasitic wasp *Diachasmimorpha juglandi*. *Journal of Chemical Ecology* 28(11): 2221-2224.

Jones WD, Nguyen T-AT, Kloss B, Lee KJ, Vosshall B. 2005. Functional conservation of an insect odorant receptor gene across 250 million years of evolution. *Current Biology* 15(4): R119-R121.

Kopp A, Barmina O, Hamilton AM, Higgins L, McIntyre LM, Jones CD. 2008. Evolution of gene expression in the Drosophila olfactory system. *Molecular Biology and Evolution* 25(6): 1081-1092.

Kyte J, Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* 157(1): 105-132.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21): 2947-2948.

Larsson MC, Domingos AI, Jones WD, Chiappe ME, Amrein H, Vosshall LB. 2004. *Or83b* encodes a broadly expressed odorant receptor essential for *Drosophila* olfaction. *Neuron* 43(5): 703-714.

Lartigue A, Campanacci V, Roussel A, Larsson AM, Jones TA, Tegoni M, Cambillau C. 2002. X-ray structure and ligand binding study of moth chemosensory protein. *Journal of Biological Chemistry* 277(35): 32094-32098.

Linn C, Feder JL, Nojima S, Dambroski HR, Berlocher SH, Roelofs W. 2003. Fruit odor discrimination and sympatric host race formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences USA* 100(20): 11490-11493.

Linn CE, Dambroski HR, Feder JL, Berlocher SH, Nojima S, Roelofs WL. 2004. Postzygotic isolating factor in sympatric speciation in *Rhagoletis* flies: Reduced response of hybrids to parental host-fruit odors. *Proceedings of the National Academy of Sciences USA* 101(51): 17753-17758.

Marck C. 1988. 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Research* 16(5 Part A): 1829-1836.

Miyazawa M, Wada T, Kameoka H. 1998. Biotransformation of (+)- and (-)-limonene by the larvae of common cutworm (*Spodoptera litura*). *Journal of Agricultural Food Chemistry* 46(1): 300-303.

Miyazawa M, Shindo M, Shimada T. 2002. Sex differences in the metabolism of (+)- and (-)-limonene enantiomers to carveol and perillyl alcohol derivatives by cytochrome P450 enzymes in rat liver microsomes.

Chemical Research in Toxicology 15(1):15-20.

Peng G, Leal WS. 2001. Identification and cloning of a pheromone-binding protein from the oriental beetle, *Exomala orientalis*. *Journal of Chemical Ecology* 27(11): 2183-2192.

Prokopy RJ, Bush GL. 1973. Ovipositional responses to different sizes of artificial fruit by flies of *Rhagoletis pomonella* species group. *Annals of the Entomological Society of America* 66: 927-929.

Ramsdell KMM. 2004. *Discovery and phylogeny of the odorant binding and chemosensory proteins of Diabrotica virgifera virgifera (Coleoptera: Chrysomelidae) and Rhagoletis (Diptera: Tephritidae)*. Ph.D. Dissertation, University of Illinois at Urbana-Champaign.

Rinehart JP, Li A, Yocum GD, Robich RM, Hayward SAL, Denlinger DL. 2007. Up-regulation of heat shock proteins is essential for cold survival during insect diapause. *Proceedings of the National Academy of Sciences USA* 104(27): 11130-11137.

Robertson HM, Martos R, Sears CR, Todres EZ, Walden KKO, Nardi JB. 1999. Diversity of odourant binding proteins revealed by an expressed sequence tag project on male *Manduca sexta* moth antennae. *Insect Molecular Biology* 8(4): 501-518.

Robertson HM, Warr CG, Carlson JR. 2003. Molecular evolution of the insect chemoreceptor superfamily in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA* 100 (Supplement 2): 14537-14542.

Rützler M, Zwiebel LJ. 2005. Molecular biology of insect olfaction: Recent progress and conceptual models. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology* 191(9): 777-90.

Souleyre EJF, Greenwood DR, Friel EN, Karunairetnam S, Newcomb RD. 2005. An alcohol acyl transferase from apple (cv. Royal Gala), MpAAT1, produces esters involved in apple fruit flavor. *FEBS Journal* 272: 3132-3144.

Vosshall LB, Amrein H, Morozov RS, Rzhetsky A, Axel R. 1999. A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell* 96:725-736.

Vosshall LB, Stocker RF. 2007. Molecular architecture of smell and taste in *Drosophila*. *Annual Review of Neuroscience* 30: 505-533.

Weinzierl RA. 1998. Botanical insecticides, soaps, and oil. In: Rechcigl JE, Rechcigl NA, editors. *Biological and Biotechnological Control of Insect Pests*. pp. 101-122. CRC Press.

Xu PX, Atkinson R, Jones DMN, Smith DP. 2005. *Drosophila* OBP LUSH is required for activity in pheromone-sensitive neurons. *Neuron* 45(2): 193-200.